BMJ Open

Incidence and clearance of oral and cervicogenital HPV infection: longitudinal analysis of the MHOC cohort study

Andrew F Brouwer ,[1] Lora P Campredon,[1] Heather M Walline,[2] Brittany M Marinelli,[2] Christine M Goudsmit,[2] Trey B Thomas,[2] Rachel L Delinger,[1] Yan Kwan Lau,[1] Emily C Andrus,[1] Thankam Nair,[2] Thomas E Carey,[2] Marisa C Eisenberg ,[1] Rafael Meza[1]

To cite: Brouwer AF, Campredon LP, Walline HM, *et al*. Incidence and clearance of oral and cervicogenital HPV infection: longitudinal analysis of the MHOC cohort study. *BMJ Open* 2022;**12**:e056502. doi:10.1136/bmjopen-2021-056502

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-056502).

MCE and RM contributed equally.

Received 16 August 2021
Accepted 09 December 2021

[1]Department of Epidemiology, University of Michigan, Ann Arbor, Michigan, USA
[2]Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan, USA

Correspondence to
Dr Andrew F Brouwer; brouweaf@umich.edu

ABSTRACT

Objectives The Michigan HPV and Oropharyngeal Cancer study aimed to evaluate patterns of oral and cervicogenital human papillomavirus (HPV) infection prevalence, incidence, and clearance as well as their relationship to sexual behaviours.

Design Cohort

Setting General public in and around Ann Arbor, Michigan.

Participants 394 college-age and older-adult participants of both sexes provided oral samples, and 325 completed at least 2 visits. 130 who provided a cervicogenital samples, and 127 completed at least 2 visits.

Outcomes Incidence and clearance rates as well as HRs for oral and cervicogenital HPV.

Results Oral HPV infections were transient, with only 16% of genotypes persisting to the next visit. The mean time to clearance of a genotype was 46 days (95% CI 37 to 58). In contrast, cervicogenital infections were more persistent, with 56% of genotypes persisting to the next visit. The mean time to clearance of a genotype was 87 days (95% CI 74 to 102). HPV vaccination was associated with reduced incidence of cervicogenital HPV infection (HR 0.63; 95% CI 0.47 to 0.83) but not oral HPV infection. Incidence of oral HPV infection was associated with 2+ recent deep kissing partners (HR 2.00; 95% CI 1.13 to 3.56). Incidence of both oral (HR: 1.70; 95% CI 1.08 to 2.68) and cervicogenital (HR 2.46; 95% CI 1.69 to 3.59) was associated with 2+ recent sexual partners.

Conclusions Detection of oral HPV was highly transient, but incidence was associated with recent deep kissing and sexual partners. Detection of cervicogenital HPV was more persistent, and incidence was positively associated with recent sexual partners and negatively associated with HPV vaccination.

Strengths and limitations of this study

- ► This study enrolled men and women and reports on both oral and cervicogenital human papillomavirus (HPV).
- ► This study's longitudinal cohort design allowed for inference of HPV dynamics.
- ► This study is limited by its comparatively small sample size and convenience sample design.

INTRODUCTION

The human papillomavirus (HPV) is the cause of virtually every cervical cancer and an increasing number and fraction of head and neck cancers.[1–8] Although vaccines are available that cover the most common cancer-causing genotypes, coverage is not complete among targeted age groups in the USA,[9] and there are oncogenic genotypes not covered by any of the available vaccines. In 2018, the US Preventive Services Task Force (USPSTF) updated its cervical cancer screening guidelines for women 21–65 to include an option of testing for high-risk HPV every 5 years, with or without cytology, in addition to the option of cervical cytology alone every 3 years.[10] While the USPSTF has concluded that the evidence for oral cancer screening in asymptomatic individuals is currently insufficient to recommend it, HPV testing could, in the future, be part of oral cancer screening either in the general population or in targeted, high-risk groups.[2] Because the most HPV infections clear without major consequences nor lead to cancer, it is essential that we understand the dynamics of cervicogenital and oral HPV infections, both to understand the implications of an oral HPV positive test and to understand the risk factors and transmission pathways associated with infection.

Cross-sectional studies, such as the National Health and Nutrition Examination Survey in the USA, can identify risk factors associated with prevalence but are unable to assess those associated with infection dynamics—neither incidence nor clearance can be determined. Longitudinal studies of HPV, such as the HPV

in Men (HIM) study,[11] have provided estimates of site-specific incidence and clearance. However, most previous longitudinal studies have had a relatively long time period between follow-up, making it difficult to understand short-term infection and clearance dynamics.

The Michigan HPV and Oropharyngeal Cancer (MHOC) study aims to evaluate patterns of oral HPV infection prevalence, incidence and clearance and their relationship to sexual history and sexual behaviours.[12] The epidemiological arm of the MHOC Study has tested a cohort of adults for oral and, in a substudy, cervicogenital HPV over 3 years, with follow-up visits every 3–4 months. This shorter follow-up time allows us to determine incidence and clearance rates in our participants with greater precision. Using a multistate transition model, we estimate the underlying rates of incidence and clearance for oral and cervicogenital HPV and the associations (HRs) of demographic and behavioural characteristics on incidence at each site.

METHODS

We previously published the full MHOC study protocol.[12] We briefly describe the main aspects of the study here.

Study subjects

Study participants were recruited in Ann Arbor, Michigan and the immediate surrounding areas. Participants were recruited at University of Michigan campus dormitories, through community fliers, and through the UM Health Research website. Volunteers over the age of 18 without a history of head and neck cancer who were willing to return every 3–4 months for 3 years for follow-up visits were invited to enrol. We enrolled 394 participants between April 2015 and December 2017. Participants completed between 1 and 12 visits, with a median of 6 visits; 325 participants completed at least two visits. A substudy focusing on cervicogenital HPV enrolled 130 participants. tudy data were collected and managed using REDCap electronic data capture tools hosted at the University of Michigan.[13 14]

Surveys

A baseline questionnaire was administered to each participant at their initial visit. Participant ID numbers were assigned to ensure participant confidentiality. Follow-up surveys were administered at each subsequent visit. The surveys were designed to individually assess a variety of topics including demographics, STI and preventive screening history, sexual health and behaviour, alcohol and drug use, and vaccination status. Vaccination status was self-reported, and due to missingness in the number of vaccine doses variable, we classified any participant reporting at least one dose of an HPV vaccine as vaccinated. Given the time frame and geographic location of the study, most vaccinated participants would have received Gardasil (6, 11, 16, 18). Sexual behaviour questions assessed current and past experiences of vaginal, oral and anal sex. The baseline questionnaire collected a complete sexual behaviour history, with the subsequent follow-up visits collecting more recent information and updates. Numbers of recent sexual partners were grouped into 0, 1, 2+ categories except for numbers of recent anal sex partners, which were grouped into 0 and 1+ because of smaller numbers.

HPV testing

All participants self-collected a saliva sample with Scope mouthwash (Procter & Gamble; Cincinnati, Ohio, USA) or an Oragene RE-100 kit (DNA Genotek; Kanata, Canada). Saliva samples were taken at each study visit. Participants who had a vagina, were not pregnant and were not menstruating at the time of a study visit were invited to self-collect a cervicogenital sample with a HerSwab (Eve Medical; Toronto, Canada). The cervicogenital substudy was rolled out after the main study, so most substudy participants had their first cervicogenital test at a follow-up visit rather than at their baseline visit. DNA was extracted from samples and genotyped using PCR Mass Array; technical details of sample processing are given in our protocol paper,[12] and technical details of the PCR Mass Array test are given in.[15] We tested for genotypes 6, 11, 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 73 and 90. Participants whose samples contained insufficient DNA or otherwise resulted in inconclusive test results were denoted as invalid.

Statistical analysis

We used Markov multistate transition modelling to estimate the incidence and clearance rate for oral HPV and cervicogenital HPV. Markov state transitions models are continuous-time, finite-state stochastic processes that assume that the transition hazard rate depends on one's current state but not on one's history (ie, we assume that previous infection does not increase the likelihood of future infection).[16] Infection and clearance occur at any time, but we only observe individuals states at certain points in time (figure 1). For a given rate of infection and clearance, we can calculate the probability of each individual's observed trajectory. By maximising this probability as a function of the infection and clearance rates, we estimate best-fit rates. Data were analysed in R V.4.0 (R Foundation for Statistical Computing; Vienna, Austria) using the msm package,[17] 2018–2020. Participants with missing data were excluded from analyses involving those missing data. Participants lost to follow-up were included if they had at least two visits.

For this analysis, we assumed that incidence and clearance of each HPV genotype occurs independently of the others and that hazard ratios are the same for all genotypes. We estimated genotype-specific rates only if there were at least 25 detections and more than one observation of persistence. We estimated HRs for incidence for selected covariates in univariable models. For these models, we assumed there is no impact of covariates on clearance—both due to the lack of biological justification

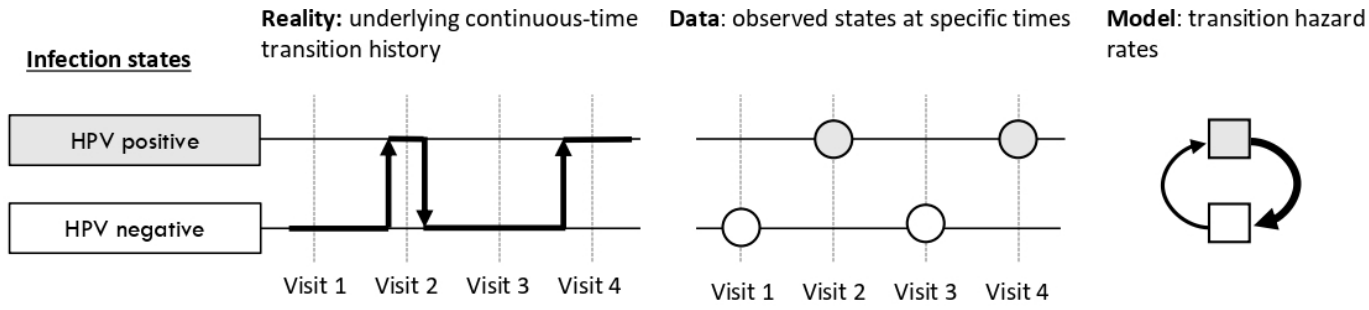

Figure 1 Participants transition between human papillomavirus (HPV) negative and positive states, and we observe these states at fixed time points. The multistate transition model estimates the underlying instantaneous infection and clearance rates that best explain the observed data when they are combined to estimate probabilities of being in each state at each visit.

for the impact of most behavioural and demographic covariates on clearance and also due to potential issues of practical unidentifiability. That is, we want to avoid estimating increased incidence as reduced clearance if we are not observing at a sufficiently fine time scale. This will potentially neglect the impact of age on clearance, but we felt that the effect of age on incidence (eg, via changes in risk, behaviour) was more salient. We also separately tested the association of the detection of multiple HPV types with clearance in a model with fixed incidence.

Patient and public involvement

Patients and the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS

Among the 325 participants who had at least two study visits, 317 had two or more valid oral HPV tests across any of their visits. The characteristics of these 317 participants are given in table 1. Oral HPV prevalence among first valid tests was 11% (34). An alluvial plot, which shows the number of participants in each state at each visit and the transition between statuses between subsequent visits, is shown in figure 2A. Among the participants, we recorded 1845 negative oral HPV tests and 148 positive oral HPV tests for at least one tested genotype. We observed 1676 pairs of participant visits: 1455 pairs of visits where the participant remained HPV negative, 94 pairs of visits where the participant transitioned from HPV negative to HPV positive, 107 pairs of visits where the participant transitioned from HPV positive to HPV negative and 20 pairs of visits in which the participant remained positive for the same genotype. (Note: the numbers of transitions will not add up to the number of tests because each participant contributes one fewer transition than their number of tests, and so the correspondence between transitions and tests depends on the specific distribution of number of tests each participant has). Only 16% of detected genotypes persisted to the next study visit. Through the multistate transition model, we estimated the average time to clearance of a previously detected genotype was 46 days (95% CI 37 to 58 days). No single genotype was detected as being persistent in an oral test more than once; accordingly, we did not estimate genotype-specific time-to-clearance for any genotypes. Time to clear one genotype was not significantly different if the participant had multiple genotypes detected (HR 1.25, 95% CI 0.65 to 2.24). Only eight individuals had multiple distinct detections of the same genotype, (ie, two positive tests with at least one negative test in between).

Among the 127 participants who provided cervicogenital samples for at least two study visits, 115 had two or more valid cervicogenital HPV tests; the characteristics of this subcohort mirror those of the full cohort, with the exception that the subcohort is entirely female. Cervicogenital HPV prevalence among first valid tests was 20% (23). The characteristics of these 115 participants are given in table 1, and alluvial plots of participant statuses are shown in figure 2B. Among these participants, we recorded 396 negative cervicogenital HPV tests and 166 positive cervicogenital HPV tests for at least one tested genotype. We observed 447 pairs of participant visits: 250 pairs of visits where the participant remained HPV negative, 74 pairs of visits where the participant transitioned from HPV negative to HPV positive, 54 pairs of visits where the participant transitioned from HPV positive to HPV negative, and 69 pairs of visits in which the participant remained positive for the same genotype. Unlike oral infections, cervicogenital infections were persistent, with 56% of detected genotypes persisting to the next study visit. Using the multistate transition model, we estimated the average time to clearance of a previously detected genotype was 87 days (95% CI 74 to 102 days). We estimated genotype-specific time-to-clearance for HPV59 (85 days, 95% CI 54 to 135), HPV66 (76 days; 95% CI 56 to 102), and HPV90 (70 days; 95% CI 47 to 104), which were all comparable. Time to clear one genotype was not significantly different if the participant had multiple genotypes detected (HR 0.79, 95% CI 0.33, 1.91). Twenty-one individuals had multiple distinct detections of the same genotype.

HRs for HPV incidence are given in table 2. In this population, participants ages 23–29 and 50+ were less likely to acquire an oral HPV infection. There were no significant differences in incidence of cervicogenital HPV by age.

Table 1 Baseline characteristics of participants in the MHOC study with at least two study visits with valid HPV tests (data collected in Ann Arbor, Michigan, USA, 2015–2017, analysed 2018–2020)

	Full cohort (N=317)		Cervicogenital substudy cohort (N=115)	
	%	n	%	n
Age				
18	29	91	25	29
19–22	33	104	32	37
23–29	12	38	11	13
30–49	12	37	16	18
50+	15	47	16	18
Sex				
Female	68	216	100	115
Male	32	101	0	0
Race				
White	60	189	64	74
Asian	23	73	18	21
Black/Hispanic/multiracial/unknown	17	55	17	20
Marital/partner status				
Never married/partnered	77	243	73	84
Ever married/partnered	23	72	27	31
Circumcised (male only)				
Yes	68	69	—	—
No	31	31	—	—
Ever diagnosed with STI*				
No	93	296	92	106
Yes	7	21	8	9
HPV vaccination				
No	45	142	45	52
Yes	48	152	50	58
Alcohol use				
Never or non-current	31	99	27	31
Current	66	210	71	82
Ever cigarette use				
Never	77	246	78	90
Ever	21	68	21	24
Ever marijuana use				
Never	54	171	53	61
Ever	41	130	44	51
Sexual attraction				
Only to another gender	72	229	73	84
Mostly to another gender	15	46	20	23
Equal or mostly/only to same gender	10	33	3	4
Deep kissing partners (6 months)				
0	42	132	79	91
1	34	109	14	16
2+	24	76	7	8
Vaginal, oral or anal sex partners (6 months)				
0	39	124	35	40
1	43	137	44	51
2+	17	54	21	24
Vaginal sex partners (6 months)				
0	49	154	43	50
1	38	120	38	44
2+	13	41	18	21
Received oral sex partners (6 months)				
0	48	152	42	48
1	36	112	39	45
2+	16	51	19	22
Performed oral sex partners (6 months)				
0	52	165	44	51
1	35	110	43	49
2+	13	40	13	15
Anal sex partners (6 months)				
0	89	279	89	101
1+	11	34	11	12

Percentages may not add up to 100% as participants could refuse to answer questions.
*Other than HPV.
HPV, human papillomavirus; MHOC, Michigan HPV and Oropharyngeal Cancer; STI, sexually transmitted infection.

Sex, race, marital status, circumcision status, previous sexually transmitted infection (STI) diagnosis, current alcohol use and ever cigarette use were not associated with incidence of either oral or cervicogenital HPV. Ever marijuana use was associated with greater incidence of cervicogenital HPV. Being vaccinated for HPV was significantly associated with lower incidence of cervicogenital HPV but not associated with incidence of oral HPV.

A greater number of deep kissing partners was associated with increased incidence of oral HPV but not significantly associated with cervicogenital HPV incidence. The number of recent (6 months) sexual partners (oral, vaginal, anal) and number of recent vaginal sex partners were each associated with greater incidence of both oral and cervicogenital HPV, with stronger associations for cervicogenital HPV. The number of recent sexual partners that one has received oral sex from or performed

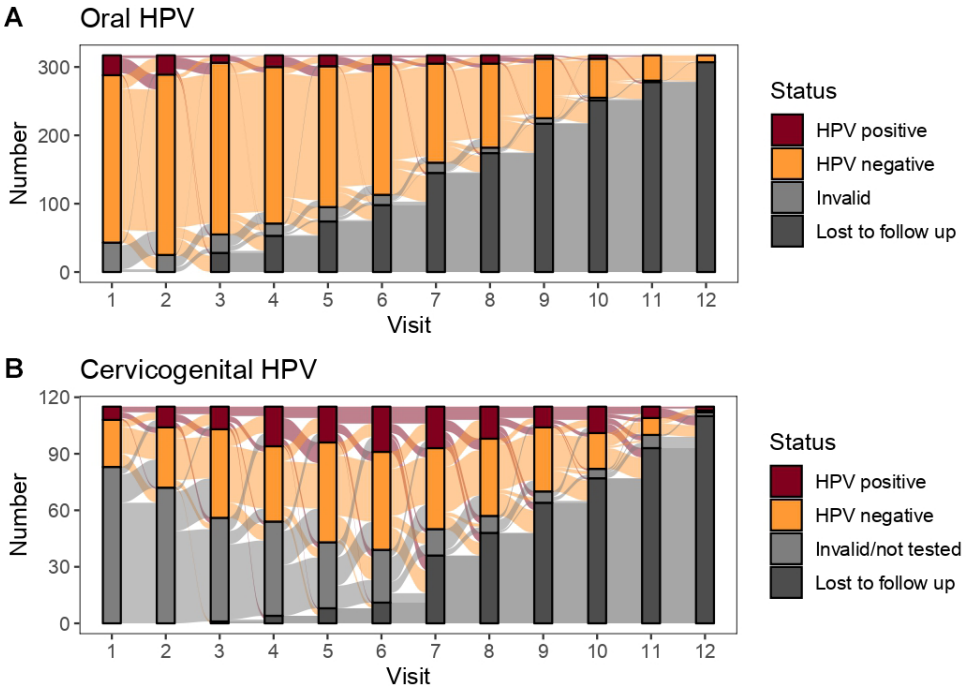

Figure 2 Alluvial plots of the longitudinal (A) oral and (B) cervicogenital human papillomavirus (HPV) status of participants in the Michigan HPV and Oropharyngeal Cancer study (data collected in Ann Arbor, Michigan, USA, 2015–2017). Note that the cervicogenital testing was rolled out later than oral testing, so that the majority of 'invalid/not tested' participants in (B) represent individuals who participated in several study visits prior to the enrolling in the cervicogenital substudy.

oral sex on were each associated with greater incidence of cervicogenital HPV but not associated with oral HPV incidence. Having at least one recent anal sex partner was not associated with either oral or cervicogenital HPV incidence.

DISCUSSION

In this study, we assessed the longitudinal dynamics of oral and cervicogenital HPV using frequent (every 3–4 months) testing over 3 years. We found that oral HPV was highly transient, with only 16% of detected genotypes persisting to the next study visit and an estimated mean of 46 days (about 1.5 months) to clearance. In contrast, cervicogenital HPV was more persistent, with 56% of detected genotypes persisting to the next study visit and an estimated mean of 87 days (about 3 months) to clearance. Incidence of oral and cervicogenital HPV were also associated with different behavioural patterns.

Previous studies estimating oral HPV clearance, including the HIM Study,[18] the Finnish Family Study[19–21] and the Persistent Oral Human Papillomavirus Study,[22] among others,[23 24] have varied substantially in their populations of interest, their sample collection and testing methodology, and their frequency of follow-up.[25 26] Estimates of time to clearance were substantially greater in the previous literature, on the order of 6 months or more, compared with the 1.5 months estimated here. Many previous studies of cervicogenital clearance, including the Hawaii Cohort Study[27] and others[28–33] have estimated mean or median clearance times of about 6–12 months, with some evidence of low-risk types clearing more quickly. In our study, we did not have the statistical power to differentiate between low-risk and high-risk genotypes, but we estimated a mean clearance time of about 3 months.

Most previous studies had comparatively long periods between follow-up, potentially obscuring underlying dynamics, particularly if clearance is fast but reinfection from a reservoir (either self or partner) is common. Other work has suggested that there may be substantial variation in short-term detectability of HPV DNA that may impact results of our and previous studies.[34] If detectability varies, then more frequent sampling is more likely to record an apparent break in infection persistence. This phenomenon could contribute to the overall shorter times to oral or cervicogenital HPV clearance in this study compared with previous studies with longer times between follow-up. We are also specifically tracking genotypes individually and not whether an individual has an infection of any HPV type, which would increase estimates of persistence. Further study of the optimal sampling frequency and methodology for oral HPV measurements is needed—if oral infection dynamics are more rapid and variable, more frequent measurements may be needed to fully assess clearance and reinfection patterns. Finally, regarding the very low persistence of oral HPV in particular, it may be that the HPV DNA we are detecting in our participants' oral cavities do not reflect true basal layer infections but rather more superficial infections. Given that PCR testing is highly sensitive and detects DNA rather than viable virions, it may also

Table 2 HRs for the incidence rate of oral and cervicogenital HPV in the MHOC study (data collected in Ann Arbor, Michigan, USA, 2015–2017, analysed 2018–2020)

	Oral HPV incidence			Cervicogenital HPV incidence		
	n	HR	95% CI	n	HR	95% CI
Age						
18	91	1 (ref)	—	29	1 (ref)	—
19–22	104	0.73	(0.49 to 1.1)	37	1.18	(0.82 to 1.69)
23–29	38	**0.32**	(0.15 to 0.68)	13	1.03	(0.63 to 1.67)
30–49	37	0.77	(0.45 to 1.29)	18	1.23	(0.78 to 1.94)
50+	47	**0.46**	(0.27 to 0.79)	18	0.92	(0.59 to 1.41)
Sex						
Female	216	1 (ref)	—	115	1 (ref)	—
Male	101	0.85	(0.59 to 1.23)	0	—	—
Race						
White	189	1 (ref)	—	74	1 (ref)	—
Asian	73	0.61	(0.37 to 1.02)	21	0.91	(0.63 to 1.32)
Black/Hispanic/multiracial/unknown	55	1.24	(0.83 to 1.85)	20	1.33	(0.95 to 1.87)
Marital/partner status						
Never married/partnered	243	1 (ref)	—	84	1 (ref)	—
Ever married/partnered	72	0.80	(0.54 to 1.19)	31	0.82	(0.59 to 1.14)
Circumcised (male only)						
Yes		1 (ref)	—	—	—	—
No		0.70	(0.33 to1.47)	—	—	—
Ever diagnosed with STI*						
No	296	1 (ref)	—	106	1 (ref)	—
Yes	21	0.81	(0.41 to 1.59)	9	1.20	(0.74 to 1.92)
HPV vaccination						
No	142	1 (ref)	—	52	1 (ref)	—
Yes	152	1.22	(0.87 to 1.71)	58	**0.63**	(0.47 to 0.83)
Alcohol use						
Never or non-current	99	1 (ref)	—	31	1 (ref)	—
Current	210	1.32	(0.91 to 1.94)	82	1.11	(0.82 to 1.51)
Ever cigarette use						
Never	246	1 (ref)	—	90	1 (ref)	—
Ever	68	1.37	(0.71 to 2.62)	24	0.92	(0.65 to 1.29)
Ever marijuana use						
Never	171	1 (ref)	—	61	1 (ref)	—
Ever	130	1.05	(0.74 to 1.47)	51	**1.48**	(1.12 to 1.96)
Sexual attraction						
Only to another gender	229	1 (ref)		84	1 (ref)	
Mostly to another gender	46	**1.57**	(1.02 to 2.43)	23	**1.53**	(1.09 to 2.17)
Equal or mostly/only to same gender	33	0.92	(0.50 to 1.68)	4	†	†
Deep kissing partners (6 months)						
0	132	1 (ref)	—		1 (ref)	—
1	109	1.65	(0.96 to 2.83)		0.87	(0.49 to 1.52)
2+	76	**2.00**	(1.13 to 3.56)		0.57	(0.25 to 1.28)

Continued

Table 2 Continued

	Oral HPV incidence			Cervicogenital HPV incidence		
	n	HR	95% CI	n	HR	95% CI
Vaginal to oral to or anal sex partners (6 months)						
0	124	1 (ref)	—		1 (ref)	—
1	137	1.26	(0.87 to 1.84)		**1.62**	(1.17 to 2.26)
2+	54	**1.70**	(1.08 to 2.68)		**2.46**	(1.69 to 3.59)
Vaginal sex partners (6 months)						
0	154	1 (ref)	—		1 (ref)	—
1	120	1.24	(0.86 to 1.78)		**1.44**	(1.05 to 1.98)
2+	41	**1.96**	(1.23 to 3.11)		**3.35**	(2.34 to 4.78)
Received oral sex partners (6 months)						
0	152	1 (ref)	—		1 (ref)	—
1	112	1.22	(0.85 to 1.74)		**1.60**	(1.18 to 2.17)
2+	51	1.07	(0.65 to 1.76)		**1.81**	(1.24 to 2.65)
Performed oral sex partners (6 months)						
0	165	1 (ref)	—		1 (ref)	—
1	110	1.41	(1.00 to 2.00)		**1.88**	(1.39 to 2.53)
2+	40	0.93	(0.52 to 1.69)		**1.97**	(1.31 to 2.97)
Anal sex partners (6 months)						
0	279	1 (ref)	—		1 (ref)	—
1+	34	0.88	(0.50 to 1.56)		1.33	(0.89 to 1.99)

Bold hazard ratios are statisitically significant at level of significance 0.05.
*Other than HPV.
†Cells with fewer than five participants are censored.
HPV, human papillomavirus; MHOC, Michigan HPV and Oropharyngeal Cancer; STI, sexually transmitted infection.

be that some of these transient detections are from non-viable virus. However, the same detection methods were used for the oral and cervicogenital samples, and we do not see the same transience in the cervicogential samples, which points to the results being driven by differences in the tissues or perhaps the collection methods.

In this analysis, HPV vaccination was associated with reduced incidence of cervicogenital HPV but not oral HPV. Previous, cross-sectional work has indicated the HPV vaccination does reduce prevalence of oral HPV.[35–37] Our longitudinal results, then, may give further credence to the hypothesis that we are detecting superficial oral infections. However, because oral HPV infections were relatively rare, we may have not had the power to detect an impact of vaccination. Cohort and age differences between our study sample and others might also explain the lack of detected association. Also, if most of the observed genotypes were not covered by the participants' vaccines (and cross-protection is likely minimal), then this result might be expected. However, of the 193 distinct detections of genotypes in oral tests, more than half (109) were type 6, 11, 16 or 18 (online supplemental table S1). In comparison, about one-fifth (36) of the 166 distinct cervicogenital detections were type 6, 11, 16 or 18. These results may suggest that vaccination had a greater impact on cervicogenital infection than on oral infection in this cohort.

Greater oral HPV incidence was associated with two or more recent deep kissing partners, vaginal sex partners, and any sex partners but was not associated with oral sex specifically. Previous literature has shown that oral HPV infection is most likely related to oral sex behaviours,[22 38 39] so our lack of association may be due to confounding. Indeed, the association between oral sex behaviour and oral HPV infection was shown to be confounded by age-cohort and race in a previous study.[39] Greater cervicogenital HPV incidence was not associated with recent deep kissing partners but was associated with one or two or more recent vaginal or oral sex partners. The number of recent sexual partners has long been known as an important risk factor for HPV, which is sexually transmitted. Ever marijuana use, which was associated with increased incidence of cervicogenital HPV infection, may not be a direct risk factor but instead be associated with true underlying risk factors that are difficult to measure directly. Although there is some laboratory evidence of immune modulation by cannabinoids,[40] epidemiological evidence for an association between marijuana use and

cervicogential HPV has been mixed,[41–44] suggesting that it is indeed likely confounded with other behaviours. Incidence of both oral and cervicogenital HPV was greater in participants who indicated sexual attraction mostly but not only to another gender; this type of 'heteroflexible' orientation has been previously associated with higher-risk sexual behaviour and STIs.[45] There was no indication of increased incidence for participants expressing sexual attraction to multiple genders equally or mostly or only to the same gender.

The strengths of this study include the longitudinal design with frequent follow-up over 3 years as well as the multistate modelling approach to assessing incidence and clearance, which enables us to use a semi-mechanistic framework to estimate covariate effects. This approach is similar to one used to analyse recurring infections in the HIM study.[46] We also use a highly sensitive PCR-based technique for HPV detection.[15] The limitations of this study include the comparatively small sample size. We are also using self-reported vaccination and behavioural data, which are subject to misclassification.

Our work contributes an additional perspective on the longitudinal dynamics of oral and cervicogenital HPV and finds substantial differences between the sites, which may have implications for the design and measurement frequency for future studies to track HPV infection and clearance dynamics. Furthermore, our infection and clearance estimates have direct application into the development of HPV transmission dynamics simulation models and of models of the natural history of HPV-related cancers.[37 47–50] Lastly, because HPV-associated cancer risk is related to persistent HPV infections, cancer screening by HPV testing requires a clear understanding of the implications of a positive HPV test. Our work emphasises that more work is needed to understand the natural history of oral HPV.

Acknowledgements We would like to thank the Michigan HPV and Oropharyngeal Cancer (M-HOC) study team for making this work possible, including Monica Yost, Jodi McCloskey, Taylor Sullivan, Alexander Mortensen, Suiyuan Huang, Keith Murphy, Bonnie Cheng, Kayla Stanek, Renata Terrazzan, Eliyas Asfaw, Mark Jung Woo Lee, Alexandra Kalabat, Ivan Montoya, Courtney Walsh, Ashley Wu, Liana Ysabel Bautista, Anna Morris, Nadine Jawad, Manila Hada, Bala Naveen Kakaraparthi, Peter Tortora, Taylor Vandenberg, Christina Hanson, Lucy Yang, Macy Afsari, Alanna Clark, Anna Gottschlich, Chinmay Pandit, Greg Foakes, Kristin Bevilacqua, Jesse Contreras, Maxwell Salvatore, Christian Alvarez, Pianpian Cao, Kelly Sun, Sheila Terrones and Miranda West.

Contributors AFB: formal analysis, methodology, software, writing-original draft, writing-review and editing. LPC: data curation, investigation, project administration, supervision, writing-review and editing. HMW: investigation, supervision, writing-review and editing. BMM: investigation, project administration. CMG: investigation. TBT: investigation. RLD: investigation, project administration. YKL: investigation, project administration. ECA: investigation, project administration. TN: investigation. TEC: conceptualisation, funding acquisition, supervision, writing-review and editing. RM: conceptualisation, funding acquisition, supervision, writing-review and editing, guarantor. MCE: conceptualisation, funding acquisition, supervision, writing-review and editing, guarantor.

Funding This work was supported by National Institutes for Health grant U01CA182915. Data management was supported by the Michigan Institute for Clinical and Health Research (CTSA grant UL1TR002240).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study was approved by University of Michigan IRB (HUM00090236).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The datasets generated and/or analysed during the current study are not publicly available because of participant privacy concerns but are available from the corresponding author on reasonable request. IRB approval or a data use agreement may be required.

ORCID iDs
Andrew F Brouwer http://orcid.org/0000-0002-3779-5287
Marisa C Eisenberg http://orcid.org/0000-0003-4221-830X

REFERENCES

1 Jemal A, Simard EP, Dorell C, *et al*. Annual Report to the Nation on the Status of Cancer, 1975-2009, featuring the burden and trends in human papillomavirus(HPV)-associated cancers and HPV vaccination coverage levels. *J Natl Cancer Inst* 2013;105:175–201.
2 Gillison ML, Alemany L, Snijders PJF, *et al*. Human papillomavirus and diseases of the upper airway: head and neck cancer and respiratory papillomatosis. *Vaccine* 2012;30 Suppl 5:F34–54.
3 Serrano B, Brotons M, Bosch FX, *et al*. Epidemiology and burden of HPV-related disease. *Best Pract Res Clin Obstet Gynaecol* 2018;47:14–26.
4 de Martel C, Plummer M, Vignat J, *et al*. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* 2017;141:664–70.
5 Chaturvedi AK, Engels EA, Anderson WF, *et al*. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol* 2008;26:612–9.
6 Chaturvedi AK, Engels EA, Pfeiffer RM, *et al*. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol* 2011;29:4294–301.
7 Chaturvedi AK, Anderson WF, Lortet-Tieulent J, *et al*. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *J Clin Oncol* 2013;31:4550–9.
8 Chaturvedi AK, D'Souza G, Gillison ML, *et al*. Burden of HPV-positive oropharynx cancers among ever and never smokers in the U.S. population. *Oral Oncol* 2016;60:61–7.
9 Elam-Evans LD, Yankey D, Singleton JA, *et al*. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2019. *MMWR Morb Mortal Wkly Rep* 2020;69:1109–16.
10 Curry SJ, Krist AH, Owens DK. Screening for cervical cancer. *JAMA* 2018;320:674.
11 Giuliano AR, Lazcano-Ponce E, Villa LL, *et al*. The human papillomavirus infection in men study: human papillomavirus prevalence and type distribution among men residing in Brazil, Mexico, and the United States. *Cancer Epidemiol Biomarkers Prev* 2008;17:2036–43.
12 Eisenberg MC, Campredon LP, Brouwer AF, *et al*. Dynamics and determinants of HPV infection: the Michigan HPV and oropharyngeal cancer (M-HOC) study. *BMJ Open* 2018;8:e021618.
13 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.

14 Harris PA, Taylor R, Minor BL, *et al*. The REDCap Consortium: building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
15 Walline HM, Komarck C, McHugh JB, *et al*. High-Risk human papillomavirus detection in oropharyngeal, nasopharyngeal, and oral cavity cancers: comparison of multiple methods. *JAMA Otolaryngol Head Neck Surg* 2013;139:1320–7.
16 Durrett R. *Essentials of stochastic processes*. Springer, 1999.
17 Jackson CH. Multi-State Models for Panel Data: The msm Package for *R*. *J Stat Softw* 2011;38:1–29.
18 Kreimer AR, Pierce Campbell CM, Lin H-Y, *et al*. Incidence and clearance of oral human papillomavirus infection in men: the him cohort study. *Lancet* 2013;382:877–87.
19 Kero K, Rautava J, Syrjänen K, *et al*. Smoking increases oral HPV persistence among men: 7-year follow-up study. *Eur J Clin Microbiol Infect Dis* 2014;33:123–33.
20 Louvanto K, Rautava J, Willberg J, *et al*. Genotype-specific incidence and clearance of human papillomavirus in oral mucosa of women: a six-year follow-up study. *PLoS One* 2013;8:e53413.
21 Rautava J, Willberg J, Louvanto K, *et al*. Prevalence, genotype distribution and persistence of human papillomavirus in oral mucosa of women: a six-year follow-up study. *PLoS One* 2012;7:e42171–7.
22 Beachler DC, Sugar EA, Margolick JB, *et al*. Risk factors for acquisition and clearance of oral human papillomavirus infection among HIV-infected and HIV-uninfected adults. *Am J Epidemiol* 2015;181:40–53.
23 Edelstein ZR, Schwartz SM, Hawes S, *et al*. Rates and determinants of oral human papillomavirus infection in young men. *Sex Transm Dis* 2012;39:860–7.
24 Pickard RKL, Xiao W, Broutian TR, *et al*. The prevalence and incidence of oral human papillomavirus infection among young men and women, aged 18-30 years. *Sex Transm Dis* 2012;39:559–66.
25 Wood ZC, Bain CJ, Smith DD, *et al*. Oral human papillomavirus infection incidence and clearance: a systematic review of the literature. *J Gen Virol* 2017;98:519–26.
26 Taylor S, Bunge E, Bakker M, *et al*. The incidence, clearance and persistence of non-cervical human papillomavirus infections: a systematic review of the literature. *BMC Infect Dis* 2016;16:293.
27 Goodman MT, Shvetsov YB, McDuffie K, *et al*. Prevalence, acquisition, and clearance of cervical human papillomavirus infection among women with normal cytology: Hawaii human papillomavirus cohort study. *Cancer Res* 2008;68:8813–24.
28 Moscicki AB, Shiboski S, Broering J, *et al*. The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. *J Pediatr* 1998;132:277–84.
29 Franco EL, Villa LL, Sobrinho JP, *et al*. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. *J Infect Dis* 1999;180:1415–23.
30 Molano M, Van den Brule A, Plummer M, *et al*. Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study. *Am J Epidemiol* 2003;158:486–94.
31 Muñoz N, Méndez F, Posso H, *et al*. Incidence, duration, and determinants of cervical human papillomavirus infection in a cohort of Colombian women with normal cytological results. *J Infect Dis* 2004;190:2077–87.
32 Moscicki A-B, Ma Y, Farhat S, *et al*. Natural history of anal human papillomavirus infection in heterosexual women and risks associated with persistence. *Clin Infect Dis* 2014;58:804–11.
33 El-Zein M, Ramanakumar AV, Naud P, *et al*. Determinants of acquisition and clearance of human papillomavirus infection in previously unexposed young women. *Sex Transm Dis* 2019;46:663–9.
34 Liu S-H, Cummings DAT, Zenilman JM, *et al*. Characterizing the temporal dynamics of human papillomavirus DNA detectability using short-interval sampling. *Cancer Epidemiol Biomarkers Prev* 2014;23:200–8.
35 Chaturvedi AK, Graubard BI, Broutian T, *et al*. Prevalence of oral HPV infection in unvaccinated men and women in the United States, 2009-2016. *JAMA* 2019;322:977.
36 Hirth JM, Chang M, Resto VA, *et al*. Prevalence of oral human papillomavirus by vaccination status among young adults (18-30years old). *Vaccine* 2017;35:3446–51.
37 Brouwer AF, Eisenberg MC, Carey TE, *et al*. Multisite HPV infections in the United States (NHANES 2003-2014): an overview and synthesis. *Prev Med* 2019;123:288–98.
38 Beachler DC, Weber KM, Margolick JB, *et al*. Risk factors for oral HPV infection among a high prevalence population of HIV-positive and at-risk HIV-negative adults. *Cancer Epidemiol Biomarkers Prev* 2012;21:122–33.
39 D'Souza G, Cullen K, Bowie J, *et al*. Differences in oral sexual behaviors by gender, age, and race explain observed differences in prevalence of oral human papillomavirus infection. *PLoS One* 2014;9:e86023.
40 Klein TW, Newton C, Larsen K, *et al*. The cannabinoid system and immune modulation. *J Leukoc Biol* 2003;74:486–96.
41 Moscicki AB, Palefsky J, Gonzales J, *et al*. Human papillomavirus infection in sexually active adolescent females: prevalence and risk factors. *Pediatr Res* 1990;28:507–13.
42 de Sanjose S, Almirall R, Lloveras B, *et al*. Cervical human papillomavirus infection in the female population in Barcelona, Spain. *Sex Transm Dis* 2003;30:788–93.
43 Phelan DF, Gange SJ, Ahdieh-Grant L, *et al*. Determinants of newly detected human papillomavirus infection in HIV-infected and HIV-uninfected injection drug using women. *Sex Transm Dis* 2009;36:149–56.
44 D'Souza G, Palefsky JM, Zhong Y, *et al*. Marijuana use is not associated with cervical human papillomavirus natural history or cervical neoplasia in HIV-seropositive or HIV-seronegative women. *Cancer Epidemiol Biomarkers Prev* 2010;19:869–72.
45 Legate N, Rogge RD. Identifying basic classes of sexual orientation with latent profile analysis: developing the multivariate sexual orientation classification system. *Arch Sex Behav* 2019;48:1403–22.
46 Ranjeva SL, Baskerville EB, Dukic V, *et al*. Recurring infection with ecologically distinct HPV types can explain high prevalence and diversity. *Proc Natl Acad Sci U S A* 2017;114:13573–8.
47 Brouwer AF, Meza R, Eisenberg MC. Transmission heterogeneity and autoinoculation in a multisite infection model of HPV. *Math Biosci* 2015;270:115–25.
48 Brouwer AF, Eisenberg MC, Meza R. Case studies of gastric, lung, and oral cancer connect etiologic agent prevalence to cancer incidence. *Cancer Res* 2018;78:3386–96.
49 Hall MT, Simms KT, Lew J-B, *et al*. The projected timeframe until cervical cancer elimination in Australia: a modelling study. *Lancet Public Health* 2019;4:1–9.
50 Ryser MD, Gravitt PE, Myers ER. Mechanistic mathematical models: an underused platform for HPV research. *Papillomavirus Res* 2017;3:46–9.